

GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OK protein - protein search, using sw model

Run on: November 20, 1999, 15:08:03 : Search time 14.81 Seconds
(without alignments)

535,776 Million cell updates/sec

Title: US-09-126-945-2

Perfect score: 1.785

Sequence: 1 MGSNSPGLSSVSPSHLLPP.....GIRKPDISORLYOVFHPPI 335

Scoring table: BLOSUM62

Searched: 188963 seqs, 23686106 residues

Database : A_Geneseq_36.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Match	Length	DB	ID	Description
1	294.5	16.5	451	1	Y01520	Chicken C-11 prote
2	290.5	16.3	478	1	Y01521	Chicken C-ery prot
3	270	15.5	454	1	R3365	Sequence of putrin
4	267.5	15.1	452	1	R4356	Human HUN-FIL-1 ge
5	262.5	14.7	581	1	M4739	Human lymphoid spe
6	240.2	13.5	371	1	M5783	Human P33-18
7	238.5	13.4	520	1	M6087	Transcription fact
8	235	13.2	520	1	R78185	Protein sequence o
9	235	13.2	470	1	R78399	Murine Ets-2 protei
10	234	13.1	370	1	M60680	Murine transcrit
11	234	12.5	555	1	M49010	Polyomavirus PEX3
12	222	12.4	462	1	M00167	ELF muticx metell
13	213.5	12.1	335	1	R5351	Adenovirus E1A-P
14	194.5	10.9	548	1	M07700	Human Ets2 regios
15	194.5	10.9	548	1	M07700	Human Ets2 regios
16	151.5	8.5	272	1	R13444	Polypeptide with D
17	151.5	8.5	272	1	R13444	Human nestin. Diag
18	99	5.5	1618	1	R60127	Human nestin prote
19	97	5.4	1575	1	M84351	Murine ubiquitin-P
20	96.5	5.4	1442	1	M53516	Human retinid dege
21	94.5	5.3	1744	1	M53516	Human retinid dege
22	94.5	5.3	1744	1	M53516	Human retinid dege
23	94.5	5.3	1744	1	M53516	Human retinid dege
24	94.5	5.3	1744	1	M53516	Human retinid dege
25	88.5	5.0	1974	1	M83911	H pylori Gp50 57
26	87.5	4.9	1711	1	M70506	Osteoclast-like pr
27	87.5	4.9	1711	1	M70507	Mutant osteoblast
28	86.5	4.8	1805	1	R27204	Rat nestin. Diag
29	86.5	4.8	1805	1	R60126	Rat nestin protei
30	86.5	4.8	1805	1	R60126	Rat nestin protei
31	86.5	4.8	1805	1	R60126	Rat nestin protei
32	86.5	4.8	1805	1	R60126	Rat nestin protei
33	86.5	4.8	1805	1	R60126	Rat nestin protei
34	86.5	4.8	1805	1	R60126	Rat nestin protei
35	86.5	4.8	1805	1	R60126	Rat nestin protei
36	86.5	4.8	1805	1	R60126	Rat nestin protei
37	86.5	4.8	1805	1	R60126	Rat nestin protei
38	86.5	4.8	1805	1	R60126	Rat nestin protei
39	86.5	4.8	1805	1	R60126	Rat nestin protei
40	86.5	4.8	1805	1	R60126	Rat nestin protei
41	86.5	4.8	1805	1	R60126	Rat nestin protei
42	86.5	4.8	1805	1	R60126	Rat nestin protei
43	86.5	4.8	1805	1	R60126	Rat nestin protei

ALIGNMENTS

RESULT	ID	Score	Match	Length	DB	ID	Description
1	Y01520	294.5	16.5	451	1	Y01520	Chicken C-11 prote
2	Y01521	290.5	16.3	478	1	Y01521	Chicken C-ery prot
3	R3365	270	15.5	454	1	R3365	Sequence of putrin
4	R4356	267.5	15.1	452	1	R4356	Human HUN-FIL-1 ge
5	M4739	262.5	14.7	581	1	M4739	Human lymphoid spe
6	M5783	240.2	13.5	371	1	M5783	Human P33-18
7	M6087	238.5	13.4	520	1	M6087	Transcription fact
8	R78185	235	13.2	520	1	R78185	Protein sequence o
9	R78399	235	13.2	470	1	R78399	Murine Ets-2 protei
10	M60680	234	13.1	370	1	M60680	Murine transcrit
11	M49010	234	12.5	555	1	M49010	Polyomavirus PEX3
12	M00167	222	12.4	462	1	M00167	ELF muticx metell
13	R5351	213.5	12.1	335	1	R5351	Adenovirus E1A-P
14	M07700	194.5	10.9	548	1	M07700	Human Ets2 regios
15	M07700	194.5	10.9	548	1	M07700	Human Ets2 regios
16	R13444	151.5	8.5	272	1	R13444	Polypeptide with D
17	R13444	151.5	8.5	272	1	R13444	Human nestin. Diag
18	R60127	99	5.5	1618	1	R60127	Human nestin prote
19	M84351	97	5.4	1575	1	M84351	Murine ubiquitin-P
20	M53516	96.5	5.4	1442	1	M53516	Human retinid dege
21	M53516	94.5	5.3	1744	1	M53516	Human retinid dege
22	M53516	94.5	5.3	1744	1	M53516	Human retinid dege
23	M53516	94.5	5.3	1744	1	M53516	Human retinid dege
24	M53516	94.5	5.3	1744	1	M53516	Human retinid dege
25	M83911	88.5	5.0	1974	1	M83911	H pylori Gp50 57
26	M70506	87.5	4.9	1711	1	M70506	Osteoclast-like pr
27	M70507	87.5	4.9	1711	1	M70507	Mutant osteoblast
28	R27204	86.5	4.8	1805	1	R27204	Rat nestin. Diag
29	R60126	86.5	4.8	1805	1	R60126	Rat nestin protei
30	R60126	86.5	4.8	1805	1	R60126	Rat nestin protei
31	R60126	86.5	4.8	1805	1	R60126	Rat nestin protei
32	R60126	86.5	4.8	1805	1	R60126	Rat nestin protei
33	R60126	86.5	4.8	1805	1	R60126	Rat nestin protei
34	R60126	86.5	4.8	1805	1	R60126	Rat nestin protei
35	R60126	86.5	4.8	1805	1	R60126	Rat nestin protei
36	R60126	86.5	4.8	1805	1	R60126	Rat nestin protei
37	R60126	86.5	4.8	1805	1	R60126	Rat nestin protei
38	R60126	86.5	4.8	1805	1	R60126	Rat nestin protei
39	R60126	86.5	4.8	1805	1	R60126	Rat nestin protei
40	R60126	86.5	4.8	1805	1	R60126	Rat nestin protei
41	R60126	86.5	4.8	1805	1	R60126	Rat nestin protei
42	R60126	86.5	4.8	1805	1	R60126	Rat nestin protei
43	R60126	86.5	4.8	1805	1	R60126	Rat nestin protei

KW Chicken, C-11 protein: cell calcification inhibiting activity.
 KW cell calcification inhibiting agent: c-erg protein; arthritis deformans;
 KW ossification: spinal column ligament.
 OS Gallus sp.
 PM 21075871.A.
 PP 23-MAY-1998. 156076.
 PR 20-JUN-1997: US-878177.
 PR 18-JUN-1997: US-878177.
 PA (CHUS) CHUGAI PHARM CO LTD.
 PA (UYPE-) UNIV PENNSYLVANIA.
 PA WPI: 99-257708/22.
 DR N-PSDB: X26552.
 CC An active protein for inhibiting cell calcification - useful for
 CC the treatment of a cell, for diagnosing arthritis
 PS deformans or ossification of spinal column ligament
 PS Claim 5: Page 9-10: 15pp: Japanese.
 CC The present sequence represents chicken c-erg protein. The specification
 CC also describes a chicken C-11 protein (Y01520) which has cell
 CC calcification inhibiting activity and a cell calcification inhibiting
 CC agent containing c-erg protein. The proteins are used for measuring the
 CC calcification of a cell, for diagnosing arthritis deformans or
 CC ossification of spinal column ligament.
 SO Sequence 478 AA.

Query Match 16.3%; Score 290.5; DB 1; Length 478;
 Best Local Similarity 26.4%; Pred. No. 4,96-18;
 Matches 96; Conservative 41; Mismatches 114; Indels 113; Gaps 13;

67 YPEDSMAAKAFGASREPP-----EERPOCPYDSQAPAGSLDV 108
 Db 41 YQSTKSPRPVPOQDMLQPPARYITAMCNPNGWGRNSPDQSAAGKMYSSDV 100
 QY 109 P---GGTLEESLEQVSWGVELMDEACKLNTLTPDMSQVSKMLMTEHQ 165
 Db 101 GNVGYMEKKI--PPNMTTNE-----RAYVPADPTLSTDHVQMLEVAKE 149
 QY 166 YRLPMGC-KAQQLACGLTAMSEORQKSP-LGADVLAML-----DIV 209
 Db 150 YGLPVDVILFLQNDKELKMTKODRLPSTYNDILSHHTLAETPLPLTSDVD 209
 QY 210 K---SAAMKEKESFGAIHYCASTS-----EESM----- 236
 Db 210 KALONSRLMARHTOGATFIPNPTSYEATORTITPDLPEYDQARRAMTSHSPHTG 269
 QY 236 -----TDEVDSCSGQPIHMOLELLKPKPHYSGR 267
 Db 270 KATQSSSTVPEKTEQRPQDLPYQLGPTSSRLANSGQ-LQMGQDLE-LLSQSN 327
 QY 268 FRLMLNKEGIFKIEDSAGVAFKMLGIRKRNPAANDKISROYKRIKRPDLSORL 327
 Db 328 CTTN-ECTNGEFTKATDPEVARRGKSKKNMNTDKLSPLATRYDKNKMTVHPESP 386
 QY 328 VYOF 331
 Db 387 MKKY 390
 RESULT 3
 R33365 standard: Protein: 454 AA.
 AD R33365
 DE 30-JUN-1993 (first entry)
 DE Sequence of putine-rich repeat (GA repeat) binding protein
 KW (GABP) subunit alpha.
 KW GA binding protein: cis-regulatory element;
 KW VP16 mediated induction.
 OS Mus musculus.
 PM M0930416-A.
 PM 13-MAR-1993: 006748.
 PP 15-AUG-1991: US-746032.

PA (CARM-) CARNegie INST WASHINGTON.
 PR LAMARCO KENNETH SL, Thompson CC;
 DR WPI: 93-09398/11.
 DR N-PSDB: 037480.
 PT DNA encoding GA binding protein sub-unit - allows investigation
 PT of sub-unit sequence motif functions, for control of rapid cell
 PT division e.g. In cancer.
 PS Disclosure: Fig 2a: 68pp; English.
 CC A cis-regulatory element regulated for virion associated protein VP16
 CC gene inductions of herpes simplex virus 1 (HSV1) immediate early
 CC (IE) gene products of the A protein repeats of the putine-rich
 CC hexanucleotide 5'-GGGAR-3'. A protein repeats of the A protein has been
 CC identified in soluble preparations of rat liver nucleus. This GA
 CC identified in soluble preparations of two separable subunits.
 CC Applicants have isolated cDNA clones encoding both subunits of GABP
 CC and have revealed that one (GABP alpha) is related to the Ets
 CC transforming protein, while the other (GABP beta) contains a
 CC sequence of 33 amino acid repeats related in sequence to a variety
 CC of proteins.
 SO Sequence 454 AA.

Query Match 15.5%; Score 277; DB 1; Length 454;
 Best Local Similarity 28.8%; Pred. No. 7,4e-17;
 Matches 83; Conservative 43; Mismatches 116; Indels 46; Gaps 9;

78 PGRSSREPPPEPPPOCPYDSQAPAGSLDVPGCLTEHSLEQVSWGVELKDIETRA 137
 Db 126 PRAHHEAEALHYERAOVI-----TIDIKHITTSISERTSDVPRMA--ALECRRE 176
 QY 138 CCLINTADPMDSPSWQKMLMTEHOYRLPMGKAFQELAGKELCAMSEORQKSP 197
 Db 177 QERGSLPTDPTKMSIDVLAHWVWVKFESNDIDLTLTNSGRELCLNODPEFORHP- 236
 QY 198 GGDVLAALQITKSAAMKE-----TFQAIHYCASTS----- 232
 Db 236 KETLSLSLELRLKRYVASQEQMNEITIDQVOIIPASVPAPVPTTIKVINSSAAK 295
 QY 232 -EBSMTDEVDSCSGQ-----PIHMOLELLKPKPHYSGRIMLNKEGIFKIEDS 284
 Db 296 VQNSPRTSGEDKSSFGKRTGNNQGLDLMQLELLTDKQARD-CISWVGDG-GEFKLNOP 353
 QY 285 AQVARRMGIRKRNPAANDKISFGYKRYKELTKRQDISQRTIYQV 332
 Db 354 ELVAKRGKORRKNPTNRYEKLRSRLATRYDGMCIKVO-GRRVYKPV 400

RESULT 4
 R44556 standard: Protein: 452 AA.
 AD R44556
 DE 26-MAY-1994 (first entry)
 DE Human HUV-FLI-1 gene product.
 DE Chromosomal translocation; chimeric; chimeric; Ewing sarcoma;
 KW EWS gene; malignant melanoma; hum-11-1; clone BM25;
 KW Human peripheral neuroectodermal tumour; human chromosome 11;
 OS Homo sapiens.
 PM M09323545-A.
 PM 25-NOV-1993.
 PR 19-MAY-1993: F00494.
 PR 20-MAY-1992: FR-006123.
 PA (CNRS) CNRS CENT NAT RECH SCI.
 PA Aurias A, Delattre O, Desmarte G, Melot T, Peter M;
 PT 1993; 15: 67-74.
 DR N-PSDB: 050644.
 PT New nucleic acid of EWS gene and its hybrid(s) - contg gene
 PT sequence involved in chromosomal translocation, also derived
 PT from Ewing sarcoma and melanoma
 PT Disclosure: Fig 7; 123pp; French.

CC The probe 1181 was used to screen a human marrow cDNA library
 CC (Clontech cat. # H1058). The clone 8025 was identified and
 CC sequenced. It represents the entire coding region together with
 CC 5'- and 3'-UTRs of the Hum-F11-1 gene.
 SQ Sequence 452 AA:

Query Match 15 1% Score 270, DB 1: Length 452:
 Best Local Similarity 25.9% Pred. No. 3, 1e-15:

Matches 103: Conservative 50; Mismatches 137; Indels 108; Gaps 17:

OY 2 GASPGSSVSPSHLLPPTVSRIGLEAAAGVQLERDMSPPATPEQG---LSA 57
 D 3 GTRKEALSVSDQSL-----PDSAYGAALPKKADMTAS--GSPDGGPKRNP 50
 OY 58 FTLSTFHLIYEDSSMAKAPASREE-----PPEEPCQYIDQAPASIDLYGCG 111
 D 51 L-----PQGEWINGPVNRYNKEEDHMHGSSPVDQSY-----SHCSKLYGCG 95
 OY 112 LT-----LEHSLEQVSNVGEVLDLSTCKLNTADPMDSSNVQKMLMTE 163
 D 96 ESNPNVNSYDENRGPPTNMTNE-----RRIYAPADPLTDBEHROMLEMAI 146
 OY 164 HQRLPMPGKA-RQELACKELCMSEEDCRORSFL-GQDVLAMH----- 207
 D 147 KEYSLEMDTSTFQNMDDKELCKNKEDLRAITLNTLVLSLHSTRESSLAAYNTS 206
 OY 207 -----DIKSAAMNKKERSP-----GAHYCAKTSSE----- 234
 D 207 HTDSSRLSVKEDPSYDSVRGANNMNSGLNKPPLGSAQOTISKNTQORPOFOYIL 266
 OY 234 SWTDEVDSCSGQPIHLWQFLKELLKPRSGRFRMLNKEKGFIEDSAOVARLNGI 293
 D 267 GPTSSRLAPGSGQ-IQMLQPTLE-LLSDASMSACIN-EGTNGSRMTDPDEYANMGE 323
 OY 294 RKNRPANNTDKLSIRIOYKKGIRKPIRDISORLYVCG 331
 D 324 RSKPNNTDKLSIRALRYVYDKNITRVH-CGRVAYKE 360

RESULT 5

M47237 standard: Protein: 561 AA.

ID M47238
 DT 22-MAY-1998 (first entry)
 DE Human lymphoid-specific transcription factor NERF-2.
 KW ets-related; human: lymphoid-specific transcription factor: NERF-2;
 RN screening assay.
 OS Homo sapiens.
 PN US5721113-A.
 PD 24-FEB-1998: 368281.
 PF 03-JAN-1995: US-368281.
 PI (HUMAN-1) HUMAN GENOME SCI INC.
 PA Kunsch CA, Libermann TN, Oetting JP:
 PI WPI: 98-168405/15.
 DR N-PSDB: V15530.
 PT NERF-1 and NERF-2 nucleic acids - encode ets-related human
 lymphoid-specific transcription factors
 PS Disclosure: Columns 27-30: 17pp. English.
 CC The present sequence is the ets-related lymphoid-specific
 transcription factor NERF-2, which can be used in screening assays
 for drugs that modulate NERF activity, and to treat patients having
 CC cell cycle effects, growth regulation and cell differentiation.
 CC development, retrovirus replication, T-cell cytokine production,
 CC cell cycle effects, growth regulation and cell differentiation.
 SQ Sequence 581 AA:

Query Match 14.7% Score 262.5; DB 1: Length 581:
 Best Local Similarity 25.2% Pred. No. 2, 1e-15:
 Matches 83: Conservative 52; Mismatches 108; Indels 97; Gaps 10:

OY 26 TDLKAAAGAVGLERD-----WSPSPATPEQGLSAFYLSFDKLY-PEDSMAA 75
 D 11 TLELSNVGNQENSDSEVSEYPAIVEVBSARLEQVAAVGLDDEYTLAQDVA--- 68
 OY 76 KAPASSREPEPEPDCPVIDS-----QAPASLDVPGGLTLEHSLQ 121
 D 68 -----EEOEVETNAETVASVSHNACTDKTITLAEMLHESPTCLDSSRSEF 119
 OY 122 VQSNVGEVLDLSTCKLNTADPMDSSNVQKMLMTEHGYRLPMPKRAFQELAG 181
 D 120 IIAAMPADVI--TEYVGVSTESSEPPDPS-----IPTSDSHEPMKK 162
 OY 182 ELCMSSEDFQNSPPLGQDVLAMHLDIKMSAMNKKERTSPGAHYCASTSEESMTSD 241
 D 163 K---VGRKKPTQOSPNSG-----SPELDICKKPRG----- 192
 OY 242 SSCSGQPIHLWQFLKELLKPRSGRFRMLNKEKGFIEDSAOVARLNGIKNRPNM 301
 D 192 ---KGNITFELDLDQNTKCPRIYTOREKGEIPLVDSKAVSKIMCKKRNKPNM 248
 OY 302 YDKLSRIQYKKGIRKPIRDISORLYVCG 331
 D 249 YETNGRALTYTQRLAKVE-GQRLVYOF 277

RESULT 6

M47237 standard: Protein: 521 AA.

ID M47237
 DT 22-MAY-1998 (first entry)
 DE Human lymphoid-specific transcription factor NERF-1.
 KW ets-related; human: lymphoid-specific transcription factor: NERF-1;
 RN screening assay.
 OS Homo sapiens.
 PN US5721113-A.
 PD 24-FEB-1998: 368281.
 PF 03-JAN-1995: US-368281.
 PI (HUMAN-1) HUMAN GENOME SCI INC.
 PA Kunsch CA, Libermann TN, Oetting JP:
 PI WPI: 98-168405/15.
 DR N-PSDB: V15529.
 PT NERF-1 and NERF-2 nucleic acids - encode ets-related human
 lymphoid-specific transcription factors
 PS Disclosure: Columns 21-24: 17pp. English.
 CC The present sequence is the ets-related lymphoid-specific
 transcription factor NERF-1, which can be used in screening assays
 for drugs that modulate NERF activity, and to treat patients having
 CC need of NERF-1, ets-related factors are involved in, e.g. cancer
 CC development, retrovirus replication, T-cell cytokine production,
 CC cell cycle effects, growth regulation and cell differentiation.
 SQ Sequence 521 AA:

Query Match 13.6% Score 242; DB 1: Length 521:
 Best Local Similarity 25.3% Pred. No. 1, 2e-13:
 Matches 66: Conservative 45; Mismatches 72; Indels 78; Gaps 8:

OY 99 QAPGSLDVPGLTLEHSS-----LEDOVSNVGE--VLKD----- 134
 D 7 EEPNTQDILLTRVENSVSHNACTDKTITLAEMLHESPTCLDSSRSEPTIHAAMP 66
 OY 134 ---ETKACLNTADPMDSSNVQKMLMTEHGYRLPMPKRAFQELACMSDE 190
 D 67 DVITVEVSTESSEPPDPS-----IPTSDSHEPMKK--VGRK 108
 OY 191 FQNSPPLGQDVLAMHLDIKMSAMNKKERTSPGAHYCASTSEESMTSDSCSGQPIH 250
 D 109 KTOOSPNSG-----SPELDICKKPRG-----KGNIT 137
 OY 251 LMQFLKELLKPRSGRFRMLNKEKGFIEDSAOVARLNGIKNRPNMNTYDKLSRI 310
 D 138 LMRFLDLDQNTKCPRIYTOREKGEIPLVDSKAVSKIMCKKRNKPNMNTYDKLSRI 197

```

OY 311 QYKKGITKRPDISORLYOOF 331
DB 198 YYYORGLAKVE-GORLYOOF 217

RESULT 7
ID M57834 standard: Protein: 371 AA.
AC M57834;
DB 21-SEP-1998 (first entry)
DE Human ERF3 protein.
KM ERF3; human; cancer; epithelial cancer; breast cancer; prostate cancer;
OS Histogenesis therapy.
PN EP-83908-AA2.
PR 06-MAY-1998.
PR 23-OCT-1997: 308450.
PR 15-NOV-1996: US-746789.
PR 31-OCT-1996: US-028791.
PR (SMRK) SMITHKLINE BEECHAM CORP.
PR (HONY) HONDA MOTOR CO. LTD.
PI Debonck C, Ko LJ, Symms MJ.
PI WPI: 98-242674/22.
DR N-PSDB: V40701.
PT New isolated polynucleotide and e.g. DNA, agonists and antagonists -
PT used to diagnose and treat cancer, e.g. prostate, breast, lung or
PT epithelial tumours
PT Thim 11: Fig 1a: 34pp: English.
CC Used to treat patients with cancer.
CC Used to treat patients with cancer.
CC polypeptide and expressing the polypeptide in vivo. ERF3 is used to
CC diagnose a disease or a susceptibility to a disease related to its
CC expression by determining a mutation in the nucleic acid sequence
CC encoding it. The presence of ERF3 in a sample derived from a host can be
CC detected. Antagonists of ERF3 are used to treat patients requiring
CC inhibition of ERF3 polypeptide. ERF3 can be used in the early diagnosis
CC of cancer, particularly epithelial cancer such as breast
CC and prostate cancer.
SQ Sequence 371 AA:

```

```

Query Match 13.5%; Score 240.5; DB 1; Length 371;
Best Local Similarity 22.4%; Pred. No. 1-13; Gaps 15;
Matches 87; Conservative 55; Mismatches 95; Indels 151;

```

```

OY 55 LSAVLYSLFMYLPEDSSNAKAPASNEPPEBEPCPYDISGAPASGIDLYPGGLTL 114
DB 7 ISNIFSVNTSMTSSSDSYLAVPAAITFAD-----DLY---LIL 44
OY 115 E--EHSLEDOV-SMVGAVYKDIETACKLITADPMDSPNYOKLLW--TEROYRLP 169
DB 45 SPMQNSLESTKASWLDG-----OPQMSKTOVLDMSTYVERKKYDAS 88
OY 170 PMGAAGQELACKTCAMSEFOR-QASPLGQDVLNHL-DIKKA---AM----- 216
DB 89 AIDRSQCDMGATLNCALDELRLVCPIL-QDILADQDLIDSSSDLSIELLEKQAS 147
OY 216 ---KETSFGAIHYCATSSSESTWDEYDS---SCS----- 246
DB 148 MAFLDALDPGPFDOGSPPFADELIDQOASPYHPSGAGAPSGSSDYTAGTSRSS 207
OY 246 -----GQP----- 249
DB 208 HSSSGSGSDVLDLPDIDGKLFPSDFDCKQKQDKRKRRKRGPKSLSEIYOCLEKSK 267
OY 249 -----THMQLKLELLKPHSGYGRFJRMKNGEYFPIEDSSAOVALMGIRKRRPAND 303
DB 268 HAPRGCHMEFIDILLHPELNEGLKKNENHGVFEFLRSFAVLQADMGOKKXNNMTE 327
OY 304 KLSHSFYKYKGIIRKRPDISORLYOOF 331
DB 328 KLSHAKYTKKLEIRND-GRLRYLTF 354

```

```

RESULT 8
ID M60677 standard: Protein: 371 AA.
AC M60677;
DB 09-NOV-1998 (first entry)
DE Transcription factor ESX protein.
KM ESX; epithelial-restricted with serine box; transcription factor;
OS Histogenesis therapy.
PN EP-83908-AA2.
PR 06-MAY-1998.
PR 23-OCT-1997: 308450.
PR 15-NOV-1996: US-746789.
PR 31-OCT-1996: US-028791.
PR (SMRK) SMITHKLINE BEECHAM CORP.
PR (HONY) HONDA MOTOR CO. LTD.
PI Debonck C, Ko LJ, Symms MJ.
PI WPI: 98-242674/22.
DR N-PSDB: V40701.
PT New isolated polynucleotide and e.g. DNA, agonists and antagonists -
PT used to diagnose and treat cancer, e.g. prostate, breast, lung or
PT epithelial tumours
PT Thim 11: Fig 1a: 34pp: English.
CC Used to treat patients with cancer.
CC Used to treat patients with cancer.
CC polypeptide and expressing the polypeptide in vivo. ERF3 is used to
CC diagnose a disease or a susceptibility to a disease related to its
CC expression by determining a mutation in the nucleic acid sequence
CC encoding it. The presence of ERF3 in a sample derived from a host can be
CC detected. Antagonists of ERF3 are used to treat patients requiring
CC inhibition of ERF3 polypeptide. ERF3 can be used in the early diagnosis
CC of cancer, particularly epithelial cancer such as breast
CC and prostate cancer.
SQ Sequence 371 AA:

```

```

Query Match 13.4%; Score 238.5; DB 1; Length 371;
Best Local Similarity 22.7%; Pred. No. 1-6-13;
Matches 88; Conservative 53; Mismatches 96; Indels 151; Gaps 15;
OY 55 LSAVLYSLFMYLPEDSSNAKAPASNEPPEBEPCPYDISGAPASGIDLYPGGLTL 114
DB 7 ISNIFSVNTSMTSSSDSYLAVPAAITFAD-----DLY---LIL 44
OY 115 E--EHSLEDOV-SMVGAVYKDIETACKLITADPMDSPNYOKLLW--TEROYRLP 169

```

DB 45 SPKMSLECTEKAMWLCG-----OPQMSKTOVLDWISYOVENKNYDAS 88
 OY 170 PGNKAFOELASKELCAMSEBOFR-QRSPGLSGDVLAHF-DIMKSA-----AMN----- 216
 DB 89 AIDSRCDMDGATLNCALCEALVLPGL-GDGLHMLRPLTSSSDLSMTLELHNG 147
 OY 216 ---KERTSGAIHRCASSTSEWSDSDVS-----SCS----- 246
 DB 148 MAPOELADGPFQDSPPAGELLDGQASPHYHSGSCADAPSPESADYVSTAGTASRSS 207
 OY 246 -----GQP----- 249
 DB 208 HNSDSDGSDVDLPDTGKLEPSDGF-ROCKKGDPMHCKRKRPRRLSKRYEWCLEGKSK 267
 OY 249 -----ILHMOGLKELLMLPHSGRFRIMLKNKGIFKLSDSAQVAKMGIRRNPMANNYD 303
 DB 268 HAPRGTLMETLINDILHPELNGMLKMNHESYFETKLASAVAMQMOCKNKNMITE 327
 OY 304 KLSRIMYOYKKGKIKKRPDISORLYOYF 331
 DB 328 KLSRAMYRYKREILERV-DGRLYYF 354
 RESULT 9
 ID R78185 standard: Protein: 520 AA.
 AD R78185.
 DT 09-FEB-1996 (first entry)
 DE Protein sequence of PEAS-beta -an ETS transcription factor.
 KW Transcription factor; Probe: reverse transcription; PCR: primer;
 KM expression vector; E.coli; Cos cell; ras; cancer cell multiplication;
 NM polyoma virus; transformation.
 FS Homo sapiens.
 FH Key Location/Qualifiers
 FT misc.difference 84 "encoded by CTC"
 FT misc.difference 126 "encoded by AAC"
 FT misc.difference 500 "note="encoded by AAC"
 FT J07145197-A.
 PN 06-JUN-1995.
 PD 06-JUN-1995.
 PE 25-NOV-1993: 295393.
 PR 25-NOV-1993: JP-295393.
 PA (EISA) EISA CO LTD.
 PS (EISA) EISA CO LTD.
 DR WP195-081769/31.
 DR N-PSB: 081769/31.
 PT ETS transcription factor activated by ras - may be used in the study
 PT of cancer cell proliferation and the proliferation of the polyoma
 PT virus
 PS Claim 1: Page 7-9: 9pp: Japanese.
 CC The amino acid sequence of the novel ETS transcription factor family
 CC member, PEAS-beta, the gene was isolated from a HeLa cell line cDNA
 CC library. The gene was followed by PCR using primers 091770-1,
 CC to produce a probe of 170-200 bp. The screening isolated the full
 CC length sequence of the transcript factor. The gene was inserted into
 CC the expression vectors pLuscript KS and pCDV1 for expression of the
 CC protein in E.coli and COS cells, respectively. The ETS transcription
 CC factor has specificity for and is activated by ras. It is useful as a
 CC reagent in studies for the elucidation of the mechanism of cancer cell
 CC proliferation of polyoma virus transformation of cells.
 SO Sequence 520 AA.
 Query Match 13.2%; Score 236; DB 1; Length 520;
 Best Local Similarity 23.7%; Pred. No. 4,36-13;
 Matches 98; Conservative 48; Mismatches 115; Indels 120; Gaps 18;
 OY 5 SGGTSSGSPSR---LILPDTVSYGRLEAAGAGVAGNDRSPSPAPROGSLAFSL 61
 DB 135 TPTPTLSTFHTQNLFPPOATLPTSGHAPAGPV--CGVGAPAPRSLSPERG----- 186

OY 62 YFDALYFEDSSMAAKAGASSNEEPPEECQCVI--DSQAPAGSLDLPQGLTLEHSL 119
 DB 186 -----POOTFAV-----PPRPHQLOMRKMKHPENDYPS-----EQRQ 219
 OY 120 EDV-----OSMYGVGLADLTETACKLNTADPRMDSVNWCKMLTTHOYLRP- 170
 DB 220 RLSESCHPFPPOPQVPCDNRBSYHNSRPETYPAPR--PQGFQETHPLEHGVBP 277
 OY 170 ---PMSKAFOELAG---KELCAMSEBOFRQSPGLSGDVLAHLDIWKMSAMNKEFRS 221
 DB 278 MPGFAPAGFQSPMGIKQEPDYCVDSFVFNKQSSYMGKVF----- 319
 OY 222 GAIRHRCASSTSEWSDSE---VDSSC-----SGQP-----ILHM 252
 DB 319 -----SSHSQFSTENDPRLYFDTCVYEPHELGKYQOFTYRNGSPFQKRSQDLM 371
 OY 253 QTELELLKLPASYSGRFRIMLKNKGIF-FTYDSDAQVAKMGIRRNPMANNYDLSRSTRO 311
 DB 372 QFLVTLTDDP-ANAFRIAMTG--KGFEEFLIEPEVARHMGIOKRNPMANNYDLSRLRY 428
 OY 312 YKRGIRRPDISORLYOYF 332
 DB 429 YERKGINQV-AGERYYKVF 448
 RESULT 10
 ID W78399 standard: Protein: 470 AA.
 AD W78399.
 DT 11-MAY-1999 (first entry)
 DE Rat Ets-2 protein.
 DE Epsilon-subunit; muscle nicotinic acetylcholine receptor; nAChR; rat;
 NM neuromuscular junction; transynaptic expression disorder; spinal cord;
 NM myasthenia gravis; familial infantile myasthenia; acetylcholinesterase;
 KM skeletal muscle tissue; Brain tissue; Alzheimer's disease;
 KM brain aneurysm; infection; tumour; bleeding; blood clot.
 OS Rattus norvegicus.
 PN W0985611-A1.
 PD 10-DEC-1998.
 PE 05-JUN-1998: 011695.
 PR 06-JUN-1997: U0-04847.
 PA (UNM) UNIV MICHIGAN.
 PS (UNM) UNIV MICHIGAN.
 DR WP199-059317/05.
 DR W0985611-A1.
 PT New isolated neuromuscular response element - used to develop products
 PT for treating e.g. neurological or muscular degenerative diseases or
 PT damage to spinal cord, peripheral nerve, skeletal muscle tissue or
 PT brain tissue
 PS Claim 32: Fig 4: 6pp: English.
 CC This sequence represents the rat Ets-2 protein while is an example of
 CC a neuronal response element (NRE) binding protein. NRE corresponds
 CC to the Ets-2 protein. The gene was isolated from a rat skeletal muscle
 CC tissue library. The gene was followed by PCR using primers 091770-1,
 CC to produce a probe of the Ets-2 protein. The screening isolated the full
 CC length sequence of the transcript factor. The gene was inserted into
 CC the expression vectors pLuscript KS and pCDV1 for expression of the
 CC protein in E.coli and COS cells, respectively. The ETS transcription
 CC factor has specificity for and is activated by ras. It is useful as a
 CC reagent in studies for the elucidation of the mechanism of cancer cell
 CC proliferation of polyoma virus transformation of cells.
 SO Sequence 470 AA.
 Query Match 13.2%; Score 236; DB 1; Length 470;
 Best Local Similarity 21.0%; Pred. No. 3,76-13;
 Matches 86; Conservative 44; Mismatches 105; Indels 174; Gaps 12;
 OY 95 VIDSOAMQASLDLYPGGLTLEHS-----LEQSVHVEYELK---DIFNCKRLM 142

```

Db 39 VLPSLSEEDTQIEVPTGLDSVSDATCEPLPLTPCSCAAVMSQALKATSGFHKQERBLG 98
Oy 143 ITADPMDSFNSVOKMLLMTERTROYLPFGPKAGFOLAEKELCMSEQSPRQSP-AGSDV 201
Db 99 IFRNMLNMQVOOQMLMNTNBSLNVNMLORFGNMQOMLCMLKREPLTADPVDGI 158
Oy 202 LKHLID-----TW-----KSNANFE-----218
Db 159 LWEHLEOMIKENOKETROYEENSHLVANPINSNTLGGFVQAPYQAPSYLIDOL 218
Oy 218 -----RTSPGAIHYCASTSE-----233
Db 219 DQMKCPSPALPALASFOQLMLPKSKRLMTAVSNVTSISQDFPGNLMNLSSGKFKENH 278
Oy 233 -----ESMTD-----SPNDSSC-----246
Db 279 SPENGDSPSSDSELLRNSNSQSLDVOVPSEFSEFECQSGSLCSLPTMSKDYIOE 338
Oy 246 -----GPIHMQLEKELMLFPHSGFRIMLMEKGIKIED 283
Db 339 RSDPEQGRPVIPAVLAGTSGPGLMQGLE-LLSDSCQSPISWTG-DQMEFKLAD 396
Oy 284 SANQVARIKGRPNMNYDKLSRTIOYKKGKIKINIKDLSQALYQY 332
Db 397 PQEVARRMCKRKRKPNMYEKLSRGLAYYDKNLIKTS-GRRYVNF 444

```

RESULT 11

R60680 standard: Protein: 370 AA.

AC M60680: 09-NOV-1998 (first entry)

DE Murine transcription factor ESX protein.

KM ESX; epithelial-restricted with serine box; transcription factor;

KW EST; mouse; epithelial cancer; breast cancer; adenocarcinoma;

MS diagnosis: therapy.

PS Mus sp.

FT Region Location/Qualifiers

FT 1..62 "first variable region"

FT Domain 63..102

FT 103..187 "A-region/poised domain"

FT Domain 188..238 "second variable region (claim 1)"

FT Misc_difference 188..238 "transactivation domain"

FT 188..238 "residue critical for transactivation"

FT Region 239..273 "serine-rich box"

FT 274..354 "third variable region"

FT Domain 355..375 "fourth variable domain"

FT Region 376..422 "fourth variable domain"

PN M09823782-A2.

PD 04-JUN-1998.

PE 26-NOV-1997; US-01865.

PR 25-NOV-1997; US-011504.

PR 23-NOV-1998; US-011504.

PR (RESC) UNIV CALIFORNIA

PR (RESC) UNIV CALIFORNIA

PR (RESC) UNIV CALIFORNIA

PR (RESC) UNIV CALIFORNIA

PR (RESC) UNIV CALIFORNIA

PR (RESC) UNIV CALIFORNIA

PR (RESC) UNIV CALIFORNIA

PR (RESC) UNIV CALIFORNIA

PR (RESC) UNIV CALIFORNIA

PR (RESC) UNIV CALIFORNIA

PR (RESC) UNIV CALIFORNIA

PR (RESC) UNIV CALIFORNIA

designated ETX for epithelial-restricted with serine box. Human ESX (see M60677) is a serine-rich protein expressed in all breast cancer, being expressed at the ductal carcinoma in situ stage. The ESX gene that results in increased expression is indicative of epithelial cancer (especially breast cancer but also gastric, ovarian and lung adenocarcinoma) and of unfavourable prognosis. Anti-ESX antibodies, antisense nucleic acids or inactive ESX mutant proteins that inhibit activity of ESX can be used to inhibit growth of epithelial cancer cells, particularly cancer. ESX polypeptides, polynucleotides, antibodies and antisense agents for diagnosis and monitoring of cancer and to screen for ESX modulators of ESX, especially potential anticancer agents. Sequence 370 AA.

Query Match 13.1%, Score 234, DB 1, Length 370; BLAST Local 83, Conservative 60, Mismatches 93, Indels 152, Gaps 14;

Matches 83, Conservative 60, Mismatches 93, Indels 152, Gaps 14;

Oy 55 LSANVSLFPMILYPEDSSMAKAGASRRPEPPQVPIYQSNAPSGSLVPGDGL 114

Db 7 ISNFFNFYNNVSSSEPTLAPAPPT-----EGEDLV-----LTL 43

Oy 115 ESHSLDQVQSVNMGVGLKDIETACKLMLITADPMDSFNSVOKMLM--TEHQRLPPLPG 172

Db 44 INQOK-----TLSEPKASVTSERQPMKTOVLNLSVQEKKIDASSID 90

Oy 173 KAPQELAEKELCMSEKOPR-QSPPLGADVIHML-DIKMSA--AMKE-----218

Db 91 FSRNCNDATLSCALBELRLVTEGP-ODOLHNOIRDLNNSDELWIELLEKQMSF 149

Oy 218 -----RTSPGAIHYCASTSE-----230

Db 150 QESGLDLPDQSGFRQGLDLDGQNASP--TCSITPGQAPSSSDVSTAGTATPOS 206

Oy 230 --TSESTVTSVSSCS-----GDP-----249

Db 207 SHASDGGSDVDLDTLSKVPDPDPTDYKEKPEGRKAKRPRKLSEYWCLEKSKK 266

Oy 249 -----THMQLEKELMLFPHSGFRIMLMEKGIKIEDSAQVLRNRPANMYD 303

Db 267 HAPGTHLMETRIIDLHLBELNGKMKRNRHGFVFLSEKVAQMGOKRKNMAYE 326

Oy 304 KLSIRIOYKKGKIKRPIQSLAYQY 331

Db 327 KLSRAFRYKREILERVD-GRRLVYKF 353

RESULT 12

R48010 standard: Protein: 555 AA.

AC M49010: 20-OCT-1998 (first entry)

DE Polyomavirus PEA3 protein.

KM Polyomavirus enhancer activator; PEA3; tumour; suppressor; inhibitor;

KW transformation; HER-2; neu promoter; metastasis; cancer.

MS M0830581-A2.

PD 16-JUL-1998.

PE 12-JAN-1997; US-780835.

PR (TEKA) UNIV TEXAS SYSTEM.

PR (TEKA) UNIV TEXAS SYSTEM.

PR (TEKA) UNIV TEXAS SYSTEM.

PR (TEKA) UNIV TEXAS SYSTEM.

PR (TEKA) UNIV TEXAS SYSTEM.

PR (TEKA) UNIV TEXAS SYSTEM.

PR (TEKA) UNIV TEXAS SYSTEM.

PR (TEKA) UNIV TEXAS SYSTEM.

PR (TEKA) UNIV TEXAS SYSTEM.

PR (TEKA) UNIV TEXAS SYSTEM.

PR (TEKA) UNIV TEXAS SYSTEM.

PR (TEKA) UNIV TEXAS SYSTEM.

SQ Sequence 155 AA:

Job time: 516 sec

Query Match

12.1%: Score 215.5; DB 1; Length 155;

Best Local Similarity 55.3%; Pred. No. 5.2e-12;

Matches 47; Conservative 14; Mismatches 19; Indels 5; Gaps 4;

QY 249 ITHQOFKELIKHSYGRFRLNKNKRGFRIEDSAQVA*KGIRKNRPANPKR 307

DB 12 LQMOFLVLLDDP-TNAHFIATG-T-RGMETKLEPEVARIKGIKNRPANPKR 68

QY 308 SIROYKKGIIRKRPDISRLYQFV 332

DB 69 SLRYEKGIMOKV-AGERVYKVV 92

RESULT 15

ID W07700 standard; Protein; 548 AA.

AC W07700:

DT 06-APR-1997 (first entry)

DB Human ETS2 repressor factor (ERF).

KM ETS2 repressor factor; ERF; transcriptional repressor;

OS Homo sapiens. tumour; cancer; oncoprotein; therapy.

FH Key Location/Qualifiers

FT domain 29..106

FT domain /label=DNA-binding domain

FT domain /note="ets-like DNA binding domain"

FT domain /label=Active_repressor_domain

FT domain /note="(claim 20)"

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain

FT domain